



MONTCLAIR STATE
UNIVERSITY

Montclair State University
**Montclair State University Digital
Commons**

Theses, Dissertations and Culminating Projects

5-2013

Talker Identification Learning and New Jersey Dialect Discrimination

Michael Patrick Apfelbaum
Montclair State University

Follow this and additional works at: <https://digitalcommons.montclair.edu/etd>



Part of the [Psychology Commons](#)

Recommended Citation

Apfelbaum, Michael Patrick, "Talker Identification Learning and New Jersey Dialect Discrimination" (2013).
Theses, Dissertations and Culminating Projects. 721.
<https://digitalcommons.montclair.edu/etd/721>

This Thesis is brought to you for free and open access by Montclair State University Digital Commons. It has been accepted for inclusion in Theses, Dissertations and Culminating Projects by an authorized administrator of Montclair State University Digital Commons. For more information, please contact digitalcommons@montclair.edu.

MONTCLAIR STATE UNIVERSITY

Talker Identification Learning and New Jersey Dialect Discrimination

by

Michael Patrick Apfelbaum

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfillment of the Requirements


For the Degree of

Master of Arts in Psychology

May 2013


College/School CHSS

Department Psychology

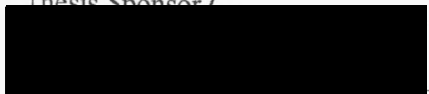

Dr. Marietta Morrissey
Dean of CHSS

April 30, 2013
(date)


Thesis Committee:


Dr. Jennifer Pardo

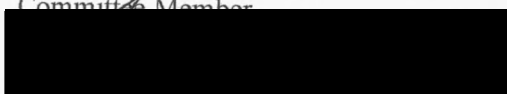
Thesis Sponsor/


Dr. Yoav Arich

Committee Member


Dr. Ruth Propper

Committee Member


Dr. Peter Vietze
Department Chair

Abstract

Previous studies have demonstrated that individuals can be trained to identify a speaker by voice. Other research has found profound regional differences in dialect across the US, with individuals being able to generally discriminate between them at above chance levels. However, little research has investigated the ability to discriminate between two similar dialects or the effect of using bisyllabic words in talker training. The present study aims to explore the patterns of talker learning and dialect discrimination which arise from training talkers on the two predominate dialects spoken in New Jersey. To investigate such factors, the current research trained 24 listeners (12 female) over several days to learn to identify speakers grouped across dialect conditions by voice. Participants were trained on a list of 80 bisyllabic words. Before and after training, listeners were tested on ability to discriminate between Northern and Southern New Jersey dialects. Results displayed a significant overall increase in ability to identify talkers over the course of training, showing a pattern of learning largely consistent with previous findings. However, no correlation between talker identification and dialect discrimination ability was found. Dialect discrimination tasks displayed no difference in score before and after training. Rates were significantly below chance levels of 50%. No evidence of a listener response bias was found, allowing the possibility that listeners were consistently discriminating between dialects but using incorrect labeling. This study fills a gap in training research in regards to bisyllabic words and provides a foundation for future research investigating discrimination between neighboring dialects.

TALKER IDENTIFICATION LEARNING AND NEW JERSEY
DIALECT DISCRIMINATION

A THESIS

Submitted in partial fulfillment of the requirements
For the degree of Master of Arts in Psychology

by

MICHAEL PATRICK APFELBAUM

Montclair State University

Montclair, NJ

2013

Copyright c 2013 by Michael Patrick Apfelbaum. All rights reserved.

ACKNOWLEDGEMENTS

My deepest gratitude is owed to Dr. Jennifer Pardo for her tremendous guidance and insight throughout the entirety of the thesis process. Her tireless assistance deeply improved not only this study, but my own research skills. Thanks to Dr. Yoav Arie and Dr. Ruth Propper for donating their time and expertise to my thesis committee. A special thanks is also in order to the individuals at the MSU Speech Communications Lab for all their co-operation and support, especially Hannah Gash for her help in running participants.

TABLE OF CONTENTS

INTRODUCTION & LITERATURE REVIEW.....	1-19
METHODS.....	19-24
RESULTS.....	25-27
DISCUSSION.....	27-34
REFERENCES.....	35-37
TABLE 1.....	38
APPENDICES.....	39-40
FIGURE CAPTIONS.....	41

Talker Identification Learning and New Jersey Dialect Discrimination

In everyday conversation, a talker gives away a world of information about themselves, revealing much more than the simple lexical meaning of their words. Speech serves not just as an outlet of inner expression, but is also replete with cues and identifiers for the external world. The human voice can be an important identifying feature of an individual, communicating information beyond its literal interpretation. Much research has been conducted on the ability to be trained to identify talkers based upon their voices. This research has shown that a general pattern of increased identification ability usually occurs as an individual completes deliberate identification training tasks (Nygaard & Pisoni, 1998). Yet our voices serve not only to distinguish us as individuals, but also to mark us as members of certain communities. This can be most easily observed by comparing individuals who speak the same language yet come from different geographic regions. Despite the relative youth of America and the common tendency to believe it is only others who speak with an accent, everyone displays aspects of a regional dialect to some extent.

Despite general similarities (Clopper & Pisoni, 2006), pronounced regional differences in dialect have been found to exist even between the residents of the geographically small state of New Jersey (Labov, 1974, 2006). This makes Montclair State University and its large student body of New Jersey residents an excellent venue from which to conduct research on dialectical differences. Though a relatively small amount of research has been performed on investigating the ability to classify dialects, results show that native speakers can generally perceive dialect across distinct regions of the United States at above chance levels (Clopper & Pisoni, 2004; Clopper, Levi, &

Pisoni, 2006). The current study will explore the pattern of learning which occurs as a listener is trained to identify a talker by voice, as well as the ability to classify between the two major dialects present in New Jersey, exploring the links between dialect identification and individual talker identification by voice.

As noted by Babel and Munson (in press), dialectical differences largely carry social connotations. An example of this often subtle phenomenon can be seen in the results of an early sociolinguistic study which found that individuals on Martha's Vineyard most threatened by tourism consistently displayed the highest levels of diphthong centralization (Labov, 1972). This diphthong centralization served as a marker of the local dialect region and was used to display the speaker's membership as a true local. Analyzing speech patterns of members of the U.S. House of Representatives, researchers found that Republican Party members were significantly more likely to pronounce the second vowel in Iraq similar to the vowel in "rack," while Democrats were more likely to pronounce using the vowel sound in "rock." This result held after controlling for factors such as regional accent, region of representation, ethnicity, sex, and age (Hall-Lew, Coppack, & Starr, 2010). Individuals in Congress were using pronunciation of an individual word as a subtle means of marking their political affiliations. Cognitive and social processes work together to select forms of pronunciation depending upon the content of language, the meaning of the message to be conveyed, and the social context. Individuals mark their identities through pronunciation style, though many individuals are not aware of the dialectical variation they produce. Overall, an individual will display membership in the larger group through consistent patterns of dialectical variation (Labov, 1974, 2006).

The state of New Jersey provides an excellent place to examine dialectal differences. Though geographically small, it is heavily populated and culturally split between two large metropolitan regions—those of New York and Philadelphia. As a result, a distinct difference in culture and language has emerged between the Southern region of the state and the densely populated Northeastern area of New Jersey. A recent study on nationwide social communication illustrated the strong divide in communication patterns between these regions. Analyzing cell phone data, researchers at IBM and MIT found that both cell phone calls and text messaging patterns split New Jersey into two separate communities (Calabrese, Dahlem, Gerber, Paul, Chen, Rowland, Rath, & Ratti., 2011). An individual in Northern New Jersey is more likely to call or text someone in Hudson Valley, New York City, or Long Island rather than someone located south or west of Ocean County, NJ. A Southerner is more likely to communicate with a friend in Eastern Pennsylvania or Delaware than an individual from the North. Thus, these communication patterns reinforce the differentiation of the two dialect patterns.

As Labov (1974, 2006) discovered in discerning dialect boundaries of Northern Pennsylvania, factors such as highways and the traffic of people between certain areas can create dialect boundaries. Considering the division of New Jersey between two major hubs of human traffic, with their own patterns of communication, the finding of distinct dialects comes as no surprise. An important outcome of these patterns of communication is the emergence of lexical differences in speech. For natives of the state, these differences can provide obvious cues as to what region a speaker comes from. In relation to gastronomy there is no difference between “sprinkles” or “jimmies,” a “sub” or a “hoagie,” “Taylor ham” or “pork roll.” However, a New Jerseyan will be quick to note

that an individual using the second of those terms likely hails from the South while one using the first of those terms comes from the North. The term “the city” will also apply to either New York or Philadelphia depending on speaker location.

These differences in language extend beyond the lexical, as the dialects of New Jersey also display important distinctions in phonology. Though the northwest of the state speaks with a less pronounced general Northern American dialect, the current study will focus on talkers from the two most distinct dialects regions in New Jersey: the New York City influenced northern region and that of the Mid-Atlantic southern region. The Northern dialect is in fact simply a part of the New York City regional dialect. This dialect is singled out by Labov (2006) as being unique in several respects. It is the smallest dialect region, geographically speaking, in the United States. The reason for this is because it is uniquely stigmatized and internalized by its speakers for purposeful correction of their speech. The New York City dialect can be most easily distinguished by features such as the frequent dropping of the letter “r” at the end of syllables, along with the intrusion of “r” in words where it does not belong, such as in “law” or “idea.” The “aw” sound is raised significantly in words such as “coffee” or “thought.” For instance, the first syllable in words such as “orange,” “horrible,” and “Florida” will have a vowel sound like that found in “card” while for most of the United States it would have a vowel sound like that of “cord.”

The dialect of Southern New Jersey is mostly similar to that of Philadelphia, often classified as Mid-Atlantic, and shares many features in common with the New York City dialect. Both display a tense-lax split of short-a vowels, meaning that while many dialects pronounce all short-a sounds in either a lax manner (the “a” in “bat”) or in a tense manner

(the “a” in “bad”), in these regions both types of short-a's exist. Both regions also raise their “aw” sounds to some extent. However, typical members of the Philadelphia region pronounce “r” at the end of words, which is an unusual feature for an east coast city. A key distinction of the Philadelphia dialect revolves around fronting back upgliding vowels, while New York City does not display any such fronting (Labov, 2006). Vowels in words such as “boot”, “bout”, or “boat” sound overemphasized to a naïve listener, comparable to production styles of a Southern dialect. There is no significant difference between the “iu” diphthong (seen in “view” or “new”) and the “u” vowel found in “boot” or “goose”. An exception to this region's fronting comes when such a vowel precedes an “r” or “l”. For instance, words with the diphthong “ow” will begin with a short-a sound (“owl” sounds like “all” and “foul” sounds like “fall”). The dialect regions and boundaries used in the current research are taken from the work of Labov (2006), which holds the New York City area dialect to be more closely related to a General Northern US dialect and distinct from the Mid-Atlantic dialect of Southern New Jersey. A map of these boundaries can be found in Appendix A.

It must be noted that a distinction between North and South Jersey is not always made in linguistic research. Clopper and Pisoni (2006) established a corpus of US dialects that was also based on Labov’s work, but their divisions were at a more macro-level. The corpus consisted of 60 native English speaking 18-25 year-olds from the University of Indiana—5 males and 5 females from 6 dialect regions recorded a variety of words, sentences, passages and interview speech. They identified 6 general dialect regions, condensing all of New Jersey and New York City into their “Mid-Atlantic” region, though they did note the unique nature of the New York City dialect. This

grouping follows from the notion that that the two dialects of New Jersey share enough important similarities that a non-native listener would have difficulty distinguishing their differences.

Prior to establishing their own corpus, Clopper and Pisoni (2006) examined naïve listener's abilities to categorize sentence-length recordings into dialect regions. Dialect regions and talkers were taken from the TIMIT corpus, representing the North, North Midland, West, New England, South Midland, and South regions. Listeners first heard 66 talkers speak the same sentence, then heard the same 66 talkers say a second sentence. After each sentence, listeners were asked to select the region they believed the talker came from. Though the overall results were poor, at around 30% correct identification, they were above 17% chance levels. Further analysis revealed that errors in responses were usually not random, and that listeners were making their judgments based on clustering 3 groups together: South and South Midland; New England with the occasional North; and North, North Midland, and West. Listeners appeared sensitive to perceptual similarities on a broader scale, using those 3 dialect groups more accurately than the 6 regions given. All of New Jersey was categorized as "New England." As researchers predicted, dropping the letter "r" played an important role in categorizing these talkers. The first sentence, containing robust examples of r-dropping was correctly identified at a higher rate than the second sentence, which did not contain such examples. As the Northern New Jersey dialect exhibits r-dropping, while the Southern New Jersey dialect does not, this may serve as an important differentiating phonetic property for listeners in the current study. Overall, researchers found that listeners relied on a smaller set of phonetic properties to differentiate between dialects than anticipated. They concluded

that categorizing an unfamiliar talker's dialect based on a short sample of speech is quite difficult for a naïve untrained listener, as listeners seldom make explicit judgments on talker origin based on such limited stimuli.

To extend these findings, Clopper, Levi, and Pisoni (2006) examined how well native English speaking undergrads could perceive dialect similarity across four of these regions—Mid-Atlantic, South, North and General American. The same region identification task as the previous study was used. In addition, a second task assessed dialect discrimination by pairing talkers. Listeners were asked how likely it was that the two talkers came from the same region on a scale of 1-7. It was found that all participants were able to identify the correct region at above chance levels, though regional differences were seen. Hearing talkers in pairs, subjects were able to correctly identify same dialect pairs at a higher rate than different dialect pairings. Clopper and Bradlow (2008) further explored dialect classification in this corpus with a free classification task between native and non-native English speakers. Though both groups showed similar classification strategies, native speakers were significantly more accurate at classifying talkers into dialect regions.

Dialects may also vary along non-geographic lines. One rarely acknowledged, yet important, individual characteristic which varies in terms of dialect is that of prestige. William Labov singled out the New York City dialect as holding a uniquely negative prestige, which many individuals purposefully try to correct. This negative prestige also helps to explain why the New York City dialect region covers such a small geographic area. Variability in dialect may also function as a marker of socioeconomic status (SES). Labov (2006) found that though it is often assumed that there exists geographic variation

across the New York City dialect area, in reality variation in dialect in this area is simply a result of the SES of the speaker. Differences in phonology are correlated to economic status rather than one's hometown or borough. The stereotypes of “Brooklynese” and “Joisey” accents, punctuated by frequent dropping of “r” and “g,” intrusive-r and highly exaggerated raising of “aw” sounds really refer to a region-wide working class dialect. Pronunciation differences across SES groups can be seen in a study by Labov (1972) which measured how often sales associates around New York City pronounced their r's when prompted to say “Fourth Floor.” The stores tested varied across the social spectrum—from the high end of *Saks*, to middle class *Macy's*, to bargain *Klein's*. Results found that employees of the high end department store pronounced r's significantly more often than the other stores, with *Klein's* showing the highest amount of r-dropping. This confirmed that increased r-dropping is characteristic of lower SES, more working-class speech.

Research pertaining to dialect classification and identification highlights how group level variability in speech can serve as a key identifying feature in discerning the larger social group which an individual comes from. However, this variability also has significance on an individual level. Dialect can serve as an important marker of personal identity which can be used to differentiate or associate oneself with larger groups. As noted by Labov (1974, 2006), when adults move to a new region they will generally accommodate their speech patterns to those of their new dialect. Listeners rapidly adapt to pronunciation characteristics of particular speakers and generalize them to others whom they presume to speak the same dialect (Kraljic, Samuel & Brenner, 2008; Pardo, 2006). Talker variability not only has a significant impact on the perceptual processing of

speech, but also affects the memory processes of the listener (Nygaard & Pisoni, 1998). Each individual comes with a unique mix of regional and social factors combined with their own idiosyncrasies, producing a unique variation in language, known as their idiolect. As noted by Babel and Munson (in press), listeners can pick up on levels of variations from the group level to the uniquely personal. Listeners have been demonstrated to be able to discern between sources of variation which are unique to speech communities (such as dialects), to individuals, and even to unusual conditions (such as having a pen in the mouth while speaking).

Learning to Identify Speakers by Voice

Nygaard and Pisoni (1998) understood the importance of the connection between individual variability in speech production and perception. Thus, they set out to study the process of perceptual learning and adaptation to individual talkers, attempting to determine how sensitivity to talker identity affects the recognition of spoken words. In three experiments, listeners were trained to identify a set of 10 voices (5 female) over a 9 day period. They were also asked to identify novel words produced not only by the talkers they were trained upon, but also by novel talkers. The first experiment trained subjects on single-word, monosyllabic utterances, while the latter two experiments employed sentence-length utterances. In Experiment 1, a set of 19 listeners completed 9 one-hour long training sessions over a 2 week period in order to learn to recognize each talker's voice and be able to match that voice to name. The first part of each session consisted of a familiarization task in which a set of 5 words from each talker would be presented to the listener to familiarize them with each talker's voice. A recognition phase followed in which listeners had to guess which talker they heard after each single word

utterance. Ten words from each talker were used and feedback was provided if incorrect. After completing two repetitions of the first two training phases, a test phase of 100 words (10 words from each of the 10 talkers presented in random order) assessed how well listeners had learned to identify the talkers each day. This test was largely identical to the recognition phase in format, though different words were used and no feedback was provided for incorrect responses. The same 100 words were used as stimuli for each of the training phases, though stimuli were reselected from the database on each day of training so that listeners never heard the same word produced by the same talker in training. Thus maximizing the number of words each listener heard from each talker.

After 9 training sessions, subjects were given a generalization test. This test included 10 words which subjects had not been trained upon from the same 10 talkers, but was otherwise identical to the test phase in format. The results showed very large differences in performance and a clear differentiation between “good” and “poor” learners. This range in scores persisted through the final day of training, during which the poorest learner only managed 29% correct while the best learner scored 98%. Follow up experiments trained subjects on sentence length utterances and found higher rates of correct identification, starting at above chance levels from the first day of training and showing continuous improvement. The current research will also attempt to train participants over several days to learn to identify talkers by voice, with a talker training format that is largely modeled after the first experiment in this study.

Another example of talker training research comes from Geiselman and Crawley (1983), who set out to understand how subjects may remember speaker voice information without apparent intent. Subjects were trained to remember the content of 24 simple past

tense sentences spoken by 24 talkers and were told their only task was to remember the sentences. After receiving instructions, some subjects heard information on the personal histories of certain talkers. Subjects were also tested not only on sentence recall, but also to their surprise, on their ability to correctly identify the talker in a paired force-choice test. Results showed that across groups there were no differences in the intentional task of sentence recognition. Attention to sentence content was relatively uniform among subjects, showing that attention was likely not shared differentially with cognitive operations such as speaker voice processing. However, a significant difference in voice identification emerged for subjects who were provided with background information of the speaker—listeners were better able to identify those talkers with personal history descriptions. Thus, with no additional stimuli, listeners did not show incidental learning of speaker identity, however the addition of personal details significantly improved incidental learning.

A follow up experiment investigated the influence of source location for a speaker on incidental memory of speaker voice. This experiment was identical to experiment one except that some subjects heard the voices playing over speakers in different areas of the room. This experiment found that while sentence recognition did not differ between participants, significant differences emerged when subjects heard the voices played from different areas of the room. These findings highlight the importance of talker identification, showing that it can be achieved without direct intent, as several processes influence memory beyond the simple meaning of what is said. With respect to the current study, it is possible that when a listener learns to identify a talker, they may do so using

dialect related variability as an aid, thereby incidentally learning something about the dialect regions present in the study.

Training Across Language Conditions

The integration of dialect discrimination and talker learning in the current study provides an opportunity to better understand how these processes may interact and overlap. While there has been a relatively small amount of research performed in this area, a study by Winters, Levi, and Pisoni (2008) investigated strategies of talker identification learning and assessed whether individuals can be trained on a speaker in one language and be able to identify him/her speaking a different language. In two experiments, they studied the extent that language familiarity affects perception of the properties of speech. First, Native English speakers were trained to identify bilingual speakers who spoke either German or English, and then tested on their ability to identify that same talker speaking the other language. Then, listeners were trained to discriminate between bilingual speakers in pairs across language condition. Subjects were trained on a database of 360 words over 4 days with familiarization, re-familiarization and recognition tasks.

Researchers found that listeners steadily improved in their ability to identify talkers in both languages. These improvements in identification largely occurred over the course of training and did not differ significantly between the two groups. They concluded that speech carries with it enough language-independent information to make identification of a familiar talker across languages possible. Because of this, listeners are able to take their knowledge of a talker's voice and generalize it across languages, enabling successful discrimination of a bilingual talker regardless of the language they

are speaking. This language-independent indexical processing was observed in German-trained listeners, who had to identify speakers in a language they did not understand. However, English trained listeners were found to employ language-dependent indexical processing. Since they could understand the language of the speakers, they relied upon their native phonological cues to aid in identification. Thus both language-dependent and language-independent information may play a role in the processing of speech depending upon context.

Another crucial factor in speech processing, in addition to language related information, is phonological awareness. Perrachione, Del Tufo, and Gabrieli (2011) underscored this point in a study of human voice recognition. They compared dyslexic and non-dyslexic participants on their ability to learn to identify voices speaking in their native language (English) or an unfamiliar language (Mandarin). Researchers focused their attention on dyslexics due to the disorder's association with deficiencies in processing native language phonology. Participants were trained and tested on voice recognition in each language condition separately, in a counterbalanced order across listeners. They found that although individuals with dyslexia could not identify English speakers at the level of controls, they could identify speakers of an unfamiliar language (Mandarin) just as well as control subjects. They concluded that this difference occurred due to the deficiency of dyslexics in the knowledge of their native phonology. The average control speaker could rely on these phonological linguistic representations to identify speakers, while a dyslexic speaker would be unable to rely upon such representations, and would only be able to use non-linguistic information to identify the talkers. This echoes the differences between foreign and native language trained listeners

observed by Winters, Levi, & Pisoni (2008). Subjects trained on English speakers could rely on their native phonology to use language-dependent cues in identification, while German trained subjects had to solely rely upon language-independent cues. Without an understanding of native phonology, dialect identification ability is limited to language-independent voice attributes only and suffers. This finding is of importance to the current research, which will only expose listeners to their native language, as it illustrates that perception of talkers in a native language relies heavily upon phonological factors. As a function of language, these phonological factors are bound to vary across dialects.

As such findings demonstrate, many easily overlooked and seemingly minor variables can contribute to differences in the comprehension and processing of language. At the same time, the results of Nygaard and Pisoni (1998), a study which serves as a model for the talker learning portion of the current research, found a large degree of inherent variance between “good” and “poor” learners. These results underscore the importance of taking individual differences in ability into consideration when analyzing patterns of learning. Such considerations are supported by studies besides Nygaard and Pisoni which have also found similarly distinct patterns of learning. Tuller, Jantzen, and Jirsa (2008) trained Native English speakers to distinguish between dental and alveolar sounds in the unfamiliar Malayam language. Subjects were found to either became more attuned to the small distinctions in acoustics or actually became less attuned. These patterns were found to be predictable based on the initial perceptual abilities of the participant. Results of Pardo et al. (submitted) also revealed large individual differences in terms of phonetic convergence.

Variability represents another important factor which can directly impact talker learning. It is thus crucial to consider how the current research could be effected by variability of training condition. A potential lesson of such effects comes from the research of Perrachione, Lee, Ha and Wong (2011). This study aimed to analyze the ability to successfully learn a foreign language phonological contrast of pitch. Researchers found success in this task to depend on individual differences in perception abilities combined with the training program design. Previous research has generally found that high stimulus variability generally increases overall learning, as it exposes listeners to a wide variety of language features and contrasts to be learned (Clopper & Pisoni, 2006). However, Perrachione et al. (2011) discovered that high variability training was only successful in increasing learning when the participant had strong perceptual abilities to begin with. Listeners who were initially stronger in discriminating performed well, while those with weak perceiving abilities were impaired by high variability and their scores suffered. Follow up analysis revealed that the aspect of training most detrimental to perceptually weak learners was the amount of trial by trial variability. Thus individual abilities and training design must be considered in interpreting results relative to any language training exercise, with care being taken to ensure trial by trial variability is not so high that it obscures listener access to training targets.

A variable which is much simpler to observe and control for, but may also play a role in the current research is that of gender. Prior research has established that gender often interacts with dialect to affect phonology, resulting in language production differences between the sexes. Labov (2006) noted these differences primarily arise though females using more prestige forms of a dialect than men, and favoring changes in

dialect- either from prestige levels down or non-prestige levels up. Labov believes females predominantly lead in the acquisition of emerging markers of dialect, however this claim has not found consensus support among researchers. Gender differences also emerge in studies of vocal accommodation, which measure if individuals adapt speaking characteristics to the style of their conversation partners. Namy, Nygaard, and Sauerteig (2002) found that when issued a task to repeat isolated words heard over headphones, females accommodated more than males and more to males than to female speakers. Even when social motives were minimized, this difference in accommodation persisted, leading researchers to conclude that accommodation may be partly due to gender related perceptual sensitivity to vocal characteristics.

There is reason to believe gender differences may be seen in the talker training portion of the current research. A study by Nygaard and Queen (2000) set out to train listeners over 3 days on both sentence length and individual word utterances of 10 talkers, with a generalization task following training. Researchers found gender differences in listener' ability to identify talkers by voice. Females successfully identified both genders of talkers, while males performed better at identifying male talkers. It is less likely that gender differences will be seen in the dialect discrimination task however. Clopper, Conrey, and Pisoni (2005) asked listeners to categorize talkers across by dialect using sentence length utterances under three presentation conditions: male, female, and mixed gender. The study employed the same talker corpus, six dialects, and format of Clopper and Pisoni (2004). Researchers found no significant difference in categorization performance across the three conditions, concluding that gender differences do not interfere with accurate categorization of regional dialect. However, analysis of errors in

classification did reveal a gender difference. Clustering patterns errors for the female talker condition were distinct from those of the male and mixed condition. Researchers attributed this difference to the possibility that at the time the corpus was recorded, Northern women were more advanced in adopting the Northern Cities Chain Shift than their male counterparts. This finding supports the gender based distinction in dialect acquisition posited by Labov.

A key motivation for this study is to examine whether learning to identify talkers by voice might lead to incidental improvement in their ability to identify dialect. A reason to believe this may occur is that focused attention on talker phonology in training may generalize to perceptual learning of dialect. With exposure, speech that is initially difficult to understand becomes more intelligible. Listeners learn about the speech they are hearing and such learning leads to improved speech perception, known as perceptual learning. Kraljic and Samuels (2006) explored how perceptual learning is generalized and whether listeners adjust their phonetic categories for a specific speaker, essentially learning their accent. Listeners heard ambiguous phonemes during a lexical decision task and were asked to categorize these sounds on a continuum, using either the voice they had been trained upon or a novel talker. They found that listeners were able to apply learning from one phoneme and speaker to new phonemes and speakers. Listeners were adjusting their perceptions of phonemic categories and generalizing their perceptual learning to new speakers and sounds.

Another example of perceptual learning of accent comes from a study by Maye, Aslin, and Tanenhaus (2005). They exposed listeners to a 20 minute story read in standard American English on Day 1 and read in synthetically accented English with

front vowels systematically lowered on Day 2. On both days, listeners completed a lexical decision task after hearing the story. Listeners more often and more quickly indicated items with lowered front vowels were words after hearing the story read in accented English. This increase was significant even for novel items, indicating that listeners had generalized their learning of the accent beyond what they had heard in the story. A follow up experiment found that this effect held even when raised front vowel items were included in the lexical decision task: the change in perceptual learning was specific to the direction of the accent. Thus, it is possible that the attention devoted to talker phonology in the training portion of the current study leads to perceptual learning which may generalize to the ability to distinguish between dialects (provided the talker set used properly illustrates differences between the two dialects being studied).

The current study integrates previous research on dialect identification and talker identification by training a set of listeners to identify talkers from North and South New Jersey. This study replicates and extends the work of Nygaard and Pisoni (1998), with listeners being trained over several days in one of three conditions (Mixed, North, or South) to learn to identify speakers by voice. Before and after training, listeners will also be tested on their ability to distinguish between Northern and Southern New Jersey dialects on a mixed set of talkers. If talker identification relies on dialect classification, or if dialect differences assist in talker identification, then there should be a relationship between dialect and talker identification. Listeners will be trained on individual, bisyllabic words produced by a set of 16 (8 female) native English speakers primarily from New Jersey who were recorded for a prior study by Pardo et al. (submitted).

Because of findings of an inherent advantage in understanding native language cues (Clopper & Bradlow, 2008), only native speakers of English will serve as participants.

Due to a relative lack of previous research in regards to both New Jersey dialects and single word utterances in talker learning, this study will be exploratory in nature. It will primarily look towards extending dialect identification tasks to single word utterances and assessing talker identification with bisyllabic words, thereby filling a gap between monosyllabic word and sentence length training conditions as used by Nygaard and Pisoni (1998). Previous research has shown that variability in the talkers set between conditions or other incidental effects of talker training may impact listener performance on the dialect discrimination task and the patterns of talker identification learning. Thus, the current study will compare a high variability training set, in which talkers come from both North and South NJ, and a low variability training set, in which talkers come from only one region. The results were analyzed to determine whether listeners could learn to identify talkers by voice with a smaller set of bisyllabic words, over fewer training sessions, and whether a relationship existed between dialect identification and talker learning.

Methods

Participants

Talkers: A total of 16 talkers were selected out of a full set of 40 talkers recorded from a previous study (Pardo et al., submitted). These talkers were Native English speakers from the Montclair State University student population and had reported normal hearing and speech. The 16 talkers were selected based on demographic features and their ability to correctly pronounce the full set of 80 bisyllabic words. Talkers were eliminated

from consideration if their recordings had multiple mispronunciations, their age was an outlier (above 30), or they had spent multiple years living outside of New Jersey or neighboring areas of New York or Pennsylvania which speak the same dialect. These talkers were then sorted based on dialect region according to Labov's (2006) classification, eliminating any talkers from the General region. The dialect region map created to sort talkers can be found in Appendix A. From the pool of classified candidates, 16 talkers evenly divided by gender and dialect region were randomly selected. Subsets of 8 talkers were created for the three talker identification training conditions (Mixed, North, and South). Due to a coding error, one female talker thought to be from the North of New Jersey was actually from Massachusetts. Talkers ranged in age from 19 to 28 years.

Listeners: A total of 24 listeners (12 female) served as participants for this study. All were recruited via email from the Montclair State University student population. All were in the psychology program at Montclair State, with 23 undergraduates and 1 graduate student. Listeners were randomly assigned into one of three training conditions. All listeners reported normal hearing and speech abilities. All of the listeners had lived in New Jersey for the majority of their lives. Their residential histories were sorted into 6 categories for data analysis purposes: Those who had lived solely in the North (n=11), the South (n=2), the General region (n=3), or those who had mixed history of residency, but lived for the longest time in the North (n=3), the South (n=2), or the General region (n=3). Ages ranged from 19 to 27 years, with a mean of 20.8. Listeners completed 8 tasks over a 2-3 week time-span to complete this study and were paid \$10 a session for their participation.

Materials

In total, 80 bisyllabic words from 16 talkers were taken from the previously recorded corpus to create the testing materials for this study. The full set of words appears in Appendix B.

Procedures

Both dialect identification tasks, as well as the Generalization task, were completed in a soundproof booth in front of a Mac Laptop running SuperLab 4.5. Talker learning training was completed in small testing rooms on PC desktops running SuperLab 4.5. For all portions of the study listeners wore Sennheiser HMDPro headphones.

Talker Words: The words which listeners were trained and tested on were recorded by talkers in a previous experiment (Pardo et al., submitted). In that study, talkers sat in a soundproof booth in front of a Mac laptop running SuperLab 4.5. Words were presented on a computer screen one at a time, and the talkers were instructed to say each word as quickly and clearly as possible. Talkers wore Sennheiser HMD280 headsets, and were recorded onto a separate iMac computer running outside of the sound booth. The recordings were then spliced into individual word files and normalized in amplitude.

In order to create the dialect identification and talker identification training task word sets, the full set of 80 words was divided into different groups that evenly distributed vowel composition and word frequency across the groups (Kucera & Francis, 1964). These words were then evenly distributed according to frequency (high, medium high, medium low, and low) into 8 word sets for use in the experiment, with 7 sets containing 8 words (with two words from each frequency category) and one set

containing 24 words (with 6 words from each frequency category), ensuring that each participant would hear all 80 words.

Dialect Identification: Listeners completed two North versus South dialect identification tasks, on the first day of the study and after receiving 5 days of talker training (see section below). These tasks assessed the effect of training on listener ability to identify talkers by dialect. Listeners sat at a table in a sound booth in front of a computer and were instructed to listen to single word, bisyllabic utterances. After hearing each word, listeners indicated whether they believed the word they had just heard was spoken by a talker from Northern or Southern New Jersey by pressing either 1 or 0 on the keyboard. Listeners, regardless of training condition, heard all 16 talkers saying the same set of 24 bisyllabic words in randomized order, for a total of 384 word trials. The set of 24 words used for this task was identical to the set of words used for the Recognition phase of training, and can be seen in Appendix B. This task was then repeated for a second block. Prior to the task, listeners were provided a dialect map of New Jersey. After completing five talker identification training sessions, listeners performed the dialect identification task again after the generalization task.

Talker Learning Training: After the first dialect identification task, the listeners performed 5 separate days of talker learning training. This portion of the study was completed in 5-9 days by all but one listener, who took 11 days to complete the training sessions. Participants were trained to identify a group of 8 (4 female) of the 16 talkers by voice. The 24 listeners were evenly divided by gender and randomly assigned into 3 conditions. Twelve listeners (6 female) were placed into a mixed condition, where they heard a mixed group of talkers from both the North and South during the training phases.

Six listeners (3 female) were placed into the North condition, where they heard only Northern talkers during the training phase. The final 6 listeners (3 female) were placed into the South condition, where they heard only Southern talkers during the training phase. Each condition had its own talker set, with the North and South conditions containing all 8 of the North or South talkers and the Mixed condition containing 4 talkers from the North and 4 talkers from the South. Listeners heard the same talker set for the entirety of training.

Each of the five days followed a format which closely resembled that of Nygaard and Pisoni (1998): a Familiarization phase, followed by a Recognition phase, finishing with a Testing phase. Across the training phases and Generalization, listeners heard all 80 words spoken by their talkers. The distribution of words across the study can be observed in Appendix B. In the Familiarization phase, listeners heard a set of 8 words spoken by each of the 8 talkers in blocked presentation. Each day of Familiarization employed a unique set of 8 words. A name and number appeared on the screen while each talker's words were played, in order for the participant to learn to associate the voice to a specific talker. At the end of a talker's word set, listeners had to hit the space bar to move on to the next talker.

The Recognition phase followed Familiarization. Here, listeners heard words one at a time while the computer screen displayed a list of all 8 talker names and numbers. Listeners had to identify which of the 8 talkers they believed said the word by pressing the corresponding 1-8 key on the keyboard. If incorrect, the computer gave them immediate feedback as to the correct talker and they had to select the correct key to move on to the next word. Listeners heard a set of 24 words spoken by the 8 talkers in random

order. This task was repeated for a second block with the identical set of words in random order, for a total of 384 word trials. The word set used here is the same as was used in the dialect identification task and the same 24 words were used for the Recognition phases of all 5 training sessions.

The Test phase followed Recognition. Like the Recognition phase, listeners used the keyboard to identify talkers from single word utterances. However, no feedback was provided for incorrect answers. Listeners heard a set of 8 novel words spoken by the 8 talkers, with a repeated second block for a total of 128 word trials. This training format was repeated for 5 days. Listeners heard the same 24 words for the Recognition task each day, while the test phase represented a novel group of 8 words each day. Familiarization words after day one were the prior day's test words. To account for order effects, the order of training days was staggered in a Latin Square design.

Generalization: After completion of the 5th day of talker learning training, listeners were scheduled to return 7-10 days later for the Generalization phase. In this portion of the study, listeners heard 8 novel words spoken by the 8 talkers they had been trained on. This task was then completed for a second block, for a total of 128 word trials. Largely identical to the Test phase, listeners were instructed to choose the key on the keyboard (1-8) which corresponded to the talker they believe they are hearing after each word. Words were presented in random order and no feedback was given. This phase, with novel words and a week-long interval since training, was done to evaluate whether listeners retained their learning from the training and could generalize this learning to novel words.

Results

Talker Training

In order to evaluate the effect of talker training on participant performance, individual trial data for the test phase blocks of each day of talker learning training were collapsed into a percentage of correct talker identification for each day (5 days plus a Generalization task). To measure growth in listener ability over the course of training, scores were converted to measures of difference from Day 1 of training. Difference scores were calculated by subtracting Day 1 performance (in percentage correct) from the performance scores for all other training days.

Across all 24 listeners, the average performance on the Test phase increased from Day 1 to Day 5 of training (Day 1 M=64.3% (9.84), Day 2 M=71.1% (13.62), Day 3 M=75.8% (11.41), Day 4 M=77.5% (12.93), Day 5 M= 77.7% (11.88)). The Generalization test, which came 7-10 days after listener's completed Day 5, continued the increase with a mean score of 79.8% (11.15).

The pattern which emerges from this data is one of steady and continuous improvement in performance from the 1st Day of training to the Generalization test. Difference scores from Training Day 1 illustrate this incremental growth. Each day of training displayed some level of increase in correct response percentage (Day 1-2: +6.7% (11.23), Day 1-3: +11.4% (9.77), Day 1-4: +13.1% (13.22), Day 1-5: +13.4% (10.38), Day 1-Gen: +15.5% (10.25)).

Figure 1 examines the interaction of listener sex and training session upon mean performance and differential scores. Listener performance appears to differ slightly across genders, being most pronounced in the higher mean difference scores in females

during training days 3-5. However, overall the patterns of growth in listener ability do not vary significantly by sex, training condition (mixed versus single dialect), or an interaction between these factors. Analysis of variance for training performance revealed that statistically significant differences only appeared across days of training ($F(5,100)=16.65, p<0.0001$). The interaction between listener sex and training session produced results of marginal significance ($F(5,100)=1.90, p=0.10$). Analysis of variance for differences in score across training days found similar results. Only training session had a significant effect upon scores ($F(4,80)=6.40, p=0.0002$), with the interaction of sex by session remaining marginal ($F(4,80)=2.36, p=0.06$).

Insert Figure 1 Here

Dialect Identification

To assess dialect identification, individual trial data for both North vs. South dialect identification tasks was collapsed into percent correct for each listener. Across all listeners, mean performance on the initial North vs. South dialect identification task was 43% (7) correct. After completing training, listeners took the North vs. South task again, with a mean score on this trial of 42% (7) correct. T-tests found no significant difference in overall scores between NvS1 and NvS2 ($t(23)=0.45, p=0.65$). The overall mean score of NvS1 was discovered to be significantly below 50% ($t(23)=-5.19, p<0.0001$), as was also true of the score of NvS2 ($t(23)=-5.57, p<0.0001$). Identification by talker region also showed roughly equivalent rates, indicated there was no listener bias to select one region over the other. Mean scores by sex and training condition varied slightly and are displayed in Table 1.

Insert Table 1 Here

These results show a pattern of largely indiscriminate scores between dialect identification tasks. The difference in overall mean score from NvS1 to NvS2 was merely -0.6%. Figure 2 demonstrates these differences on an individual level. No consensus pattern emerges other than the differences apparent by sex. Female listeners (n=12) averaged a -2.3% drop in total score from NvS1 to NvS2 while Male listeners (n=12) averaged a 0.9% increase in total score from NvS1 to NvS2. Training condition had no significant effect upon NvS2 scores. Figure 3 illustrates the difference in correct identification from NvS1 to NvS2 at the level of individual words. No consistent pattern of differences in identification appeared across both dialect identification tasks.

Insert Figures 2 and 3 Here

Statistical analysis did not yield many significant differences in regards to this portion of the study. No effect of training on North versus South dialect discrimination task performance was found. The most compelling finding was that both dialect discrimination tasks scored at significantly below 50%, this introduces the possibility that listeners were indeed consistently discriminating between dialects, but were doing so by using the wrong labels.

Discussion

This study was influenced by previous findings of talker learning training and native language discrimination abilities. It set out to explore the patterns of talker learning and dialect discrimination which arise from training talkers on the two predominate dialects spoken in New Jersey, a small state with distinct communication patterns. To investigate such factors, 24 listeners (12 female) were trained over 5 days to learn to identify speakers from two dialect regions by voice. Participants were trained on

a list of 80 bisyllabic words spoken by a subset of 8 talkers in either Mixed, Northern, or Southern New Jersey dialect conditions. Before and after training, listeners were tested on their ability to distinguish between Northern and Southern New Jersey dialects. While this study explored a possible relationship between the ability to identify speakers by voice and the ability to identify dialects, the results of the two dialect identification tasks failed to establish the existence of such a link. No change in dialect discrimination ability was observed as a result of training, and listeners correctly identified a talker's dialect at overall levels significantly below chance. In many cases, individuals scored worse on dialect discrimination after training, and no correlation was found between post-training dialect discrimination and post-training talker Generalization scores. However, talker learning training did provide listeners with a significant overall growth in the ability to identify talkers, and this growth was retained in generalization scores.

With regard to the talker training portion of the current study, the results showed an overall pattern of steady and continuous increase in identification ability over the 5 days of training. Scores persisted into the generalization task, 7-10 after the last training session. This increase over the course of training was found to be statistically significant. Such an increase in ability to identify a talker by voice is generally consistent with the results of other studies which have utilized talker training. However, a closer examination of the data reveal several differences. The research of Nygaard and Pisoni (1998) found a distinct grouping between “good” and “poor” learners (determined by ability to score over 70% in identification after 9 days of training) in their first training related experiment. The current study did not find such a large distinction, with only 4 out of 24 listeners scored below the 70% threshold after the 5th day of training. The higher rates of

identification found in the current research can be attributed to the fact that Nygaard and Pisoni trained listeners on monosyllabic utterances, while this study utilized bisyllabic utterances. An extra syllable per word will likely provide a listener with more phonological and intonational information, thus increasing identification score. When Nygaard and Pisoni conducted further experiments using sentence length utterances, overall performance increased and no split of good and poor learners was apparent, consistent with the current findings. Though the training persisted for 4 more days than the current research, this difference is unlikely to affect results, as they found that the performance levels reached a plateau around the 5th day of training and did not fluctuate significantly afterward. The interaction between listener sex and training session, approaching marginal significance is consistent with previous research that has found sex differences to emerge across talker training (Nygaard & Queen, 2000).

Another distinction in regards to talker training research is that the current research employed a relatively small database of bisyllabic words. Other studies used from 100 (Nygaard & Pisoni, 1998) to 360 monosyllabic words (Winters, Levi & Pisoni, 2008), while others trained listeners on sentences (Geiselman & Crowley, 1983; Perrachione, Del Tufro, Gabrieli, 2011). The current research used only 80 bisyllabic words due to corpus limitations and spread out these stimuli across training in a manner distinct from previous research. Subjects were trained on the same set of 24 Recognition words plus a novel set of 8 Familiarization words spoken by all talkers each day, with remaining words reserved for the Test and Generalization phases. Nygaard and Pisoni (1998) reselected their stimuli from their 100 word database each day, to maximize the number of different words heard from each talker. The ten words produced by one talker

on the 1st day, would be produced by different talker on the 2nd day. Unlike the current research, listeners did not experience the same blocks of words being spoken by each talker on a given day or hear the same word produced by the same talker in training. As training scores were high with generalization scores showing an overall increase, it is unlikely that the size of the word database was too small to allow talker learning or had a significant impact on findings. Future talker training studies can therefore rely on a smaller set of bisyllabic words than had been previously employed.

With respect to the dialect discrimination portion of the current study, results found no significant difference in scores of mean performance before (43%) and after (42%) training. No differences were found to exist between listener conditions of Mixed, North, and South New Jersey groupings. The only significant finding to arise from this portion of the study was that scores of mean performance were found to be below 50% chance levels. Though such scores appear initially unremarkable, when the data were split into identification of North talkers versus South talkers, performance levels were the same as when they were combined. Such a finding indicates that low performance levels were not due to a bias to respond to all talkers as North or South. This shows that listeners may have actually been consistently differentiating between the North and South dialects, but were doing so by using the incorrect labels—selecting “North” for Southern talkers and “South” for Northern talkers.

Other studies involving dialect discrimination task can help explain the seemingly underwhelming results of the dialect identification task in the current study. Of the relatively small amount of US dialect discrimination research that has been performed, the majority of studies have looked at large scale differences over multiple dialect regions

across the country. Clopper and Pisoni (2006) had subjects categorize talkers into 6 dialect regions, covering the vast majority of the continental US, while a follow-up study compressed the same area into only 4 regions (Clopper, Levi, & Pisoni, 2006). The results of the first study were low (30% correct), but above chance. Crucially, further analysis revealed that mistaken labeling followed a pattern and that subjects tended to cluster the 6 regions into 3 more general areas made up of similar dialect regions. Thus even though subjects were incorrect in labeling, important differentiation in dialect was nevertheless occurring. These studies also employed sentence-length utterances, rather than single bisyllabic words. This difference could also help explain the low rates of correct identification in the current research, as full sentences provide a larger potential bank of phonological information which can be used to differentiate between dialects.

The current study provides a stark contrast with previous research by using only two neighboring, fairly similar, dialect regions. As neighbors, it is inevitable that the Mid-Atlantic and New York City dialect areas will share some important common features, such as a split short-a system and a raised “ah” sound (Labov, 2006). Despite many differences, their proximity may provide less opportunities for distinction than between two regions of greater geographical distance. In addition, proximity could also mean that listeners (who all hailed from New Jersey) have a good deal more familiarity with and exposure to both dialects than they would have between even two dialects of varied regions. This would be especially true of any listeners from the South who now attend a university in the Northern region. A listener who has lived in an area near the middle of the state may experience overlap between the two dialects. These factors of similarity, familiarity, and overlap could help to explain how listeners seemingly

displayed an ability to distinguish between the two dialects, yet consistently used incorrect labeling.

As with any research, there are several limitations, as well as potential alterations for future research which could be made to provide greater insight into the current findings. The most obvious limitation to this research was the corpus of talkers which were used. Using an existing speaker corpus meant that only per-recorded stimuli of single word utterances could be used for training and discrimination purposes. While this caused no issue in regards to training or demographic necessities, it is likely that using sentence-length utterances in dialect discrimination tasks could have helped listeners highlight differences otherwise obscured by single word utterances. Future research following a similar model could evaluate this by having half of the listeners tested on discriminating dialect with single word utterances, with the remainder hearing sentence-length utterances.

Furthermore, this study was conducted on a University campus located in one of the dialect areas being studied. Both talkers and listeners alike resided in this region whether they were native to it or not. This provides a potential minor imbalance between both talker and listener groups. Ideally all subjects would have an equal amount of familiarity with both dialects studied and recordings of talkers could have been done across the state of New Jersey. Due to the recruitment area of this study, listener populations were inherently uneven—with only a handful of listeners having lived in the Southern New Jersey region. While this did not affect the goals of this study, having an equal share of Northern and Southern listeners could have enabled another point of group comparison, providing further insight into potential differences in dialect discrimination

between regions.

The lack of previous research examining identification of similar dialect regions also provides a limitation in the generalization of any findings. Previous studies have generally not examined dialect identification between only two regions on the level of individual words. Thus, while findings have revealed difficulties in distinguishing between similar dialects, it cannot be determined whether low rates of identification seen in the current study were due to dialect similarity or single words as stimuli. Future research could use the same format but use two more distinct dialects which are relatively unfamiliar to listeners, or between two distinct dialect regions which will highlight the contrasts between familiar and unfamiliar dialect. Further analysis could also recruit listeners from outside the New York City and Mid-Atlantic regions to see if they perform at significantly worse levels than natives or display consistent discrimination. This could help to determine whether the listeners in the current research simply were indeed discriminating between dialects but using consistently incorrect labels. A potential way to have avoided this uncertainty would have been to add the label of the talker's region to the training regimen along with their name and id number. This would allow for a more direct connection between talker training and dialect discrimination to be established in the minds of listeners. While the current research never explicitly associated these factors to listeners, seeking an incidental improvement but finding none, it remains to be seen whether more intentional training could have effected dialect discrimination ability. However, such a modification would require a separate subset of talkers to be used on dialect discrimination tasks.

Despite these limitations the current study made important contributions in exploring under-analyzed areas of language and learning research. The talker identification findings filled in a gap between monosyllabic word and sentence length utterance training paradigms, showing consistency with previous results and contributing to a greater understanding of the patterns of learning that occur when trained to identify a talker by voice. Though this study did not find a connection between the ability to identify talkers by voice and dialect identification, this does not mean that there is no such connection. Future research could investigate such a potential link between more distinct, non-neighboring dialects than those employed in the current study. This study was one of the first of its kind to explore such a link and as such it provides a foundation for future studies that investigate the relationship between talker identity and dialectal variation. Though dialect identification rates were low, they were not due to response bias. This finding of potential discrimination with consistently wrong labeling demonstrates that even though an individual may not be correct in their attributions, they can pick up on enough phonological cues in single bisyllabic words to differentiate between two similar dialects. Such a discovery further illustrates the complexity with which the human mind perceives and reacts to language, as well as confirms the distinction in speech patterns between Northern and Southern New Jersey.

References

- Babel, M.E., & Munson, B. (In press). Producing socially meaningful linguistic variation. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds.), *Oxford Handbook in Speech Production*.
- Calabrese, F., Dahlem, D., Gerber, A., DeDe, P., Xiaoji, C., Rowland, J., Rath, C., & Ratti, C. (2011). *The connected states of America: Quantifying social radii of influence*. Unpublished paper presented at 2011 IEEE International Conference on Social Computing.
- Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech*, 51, 175-198.
- Clopper, C. G., Conrey, B. L., & Pisoni, D. B. (2005). Effects of talker gender on dialect categorization. *Journal of Language and Social Psychology*, 24, 182-206.
- Clopper, C. G., Levi, S. V., & Pisoni, D. B. (2006). Perceptual similarity of regional varieties of American English. *Journal of the Acoustical Society of America*, 119, 566-574.
- Clopper, C. G., & Pisoni, D. B. (2006). The nationwide speech project: A new corpus of American English dialects. *Speech Communication*, 48, 633-644.
- Geiselman, R.E., Crawley, J.E. (1983). Incidental processing of speaker characteristics: Voice as connotative information. *Journal of Verbal Language & Verbal Behavior*, 22(1), 15-23.
- Hall-Lew, L., Starr, R.L., & Coppock, E. (2012). Style-Shifting in the U.S. Congress: The vowels of 'Iraq(i)'. In Campoy, J.M. & Espinosa, J.A. (Eds.), *Style-Shifting in Public: New Perspectives on Stylistic Variation*. (45-63). Amsterdam, Netherlands:

John Benjamins Publishing.

Kraljic, T., & Samuel, A.G. (2006). Generalization in perceptual learning for speech.

Psychonomic Bulletin and Review, 13(2), 262-268.

Labov, W. (1972). The recent history of some dialect markers on the island of Martha's

Vineyard, Mass. In Davis, L.M. (Ed.), *Studies in Linguistics in Honor of Raven I.*

McDavid Jr. (81-121). Tuscaloosa, Alabama: University of Alabama Press.

Labov, W. (1972). The social stratification of (r) in New York City department stores. In

Labov, W. (Ed.), *Sociolinguistic Patterns*. (43-54). Philadelphia, Pennsylvania:

University of Pennsylvania Press.

Labov, W. (1974). Linguistic change as a form of communication. In A. Silverstein (Ed.),

Human Communication: Theoretical Explorations, 221-256.

Labov, W. (2006). New York City and the Mid-Atlantic states. In Labov, W., Ash, S., &

Boberg, C. (Eds.), *Phonological Atlas of North American English*, (233-238).

Berlin, Germany: De Gruyter Mouton Publishing.

Maye, J., Aslin, R.N., & Tanenhaus, M.K. (2008). The wecked wetch of the wast:

Lexical adaption to a novel accent. *Cognitive Science*, 32, 543-562.

Namy, L.L., Nygaard, L.C., & Sauerteig, D. (2002). Gender differences in vocal

accommodation: The role of perception. *Journal of Language and Social*

Psychology, 21(4), 422-432.

Nygaard, L.C., & Pisoni, D.B. (1998). Talker-specific learning in speech perception.

Perception and Psychophysics, 60(3), 355-376.

Nygaard, L.C., & Queen, J.S. (2000). The role of sentential prosody in learning voices.

Journal of the Acoustical Society of America, 107(5), 2856-2856.

- Pardo, J.S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119 (4), 2382-2393.
- Pardo, J.S., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. (submitted).
Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. Revision requested from *Journal of Memory & Language*.
- Perrachione, T.K., Lee, J., Ha, L.Y.Y., & Wong, P.C.M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *Journal of the Acoustical Society of America*, 130(1), 461-472.
- Perrachione, T.K., Del Tufo, S.N., & Gabrieli, J.D.E. (2011). Human voice recognition depends on language ability. *Science*, 333, 595.
- Tuller, B., Jantzen, M.G., & Jirsa, V.K. (2008). A dynamical approach to speech categorization: Two routes to learning. *New Ideas in Psychology*, 26, 208-226.
- Winters, S. J., Levi, S.V., & Pisoni, D. B. (2008). Identification and discrimination of bilingual talkers across languages. *Journal of the Acoustic Society of America*, 123(6), 4524-4538.

Table 1. Displays mean scores on both dialect discrimination tests by talker learning condition and listener gender. Standard deviations are provided in parentheses. Mixed condition listeners displayed slightly higher rates of identification, though not significantly so.

		NvS1	NvS2
Condition	Mixed (n=12)	44.1% (8.18)	44.4% (7.81)
	North (n=6)	39.4% (3.04)	39.6% (3.77)
	South (n=6)	43.9% (5.80)	40.6% (6.45)
Gender	Female (n=12)	44% (7.58)	41.6% (8.52)
	Male (n=12)	41.8% (5.90)	42.9% (4.81)

Appendix A: Dialect Map of New Jersey with Talker Hometowns



Appendix B: Word Lists

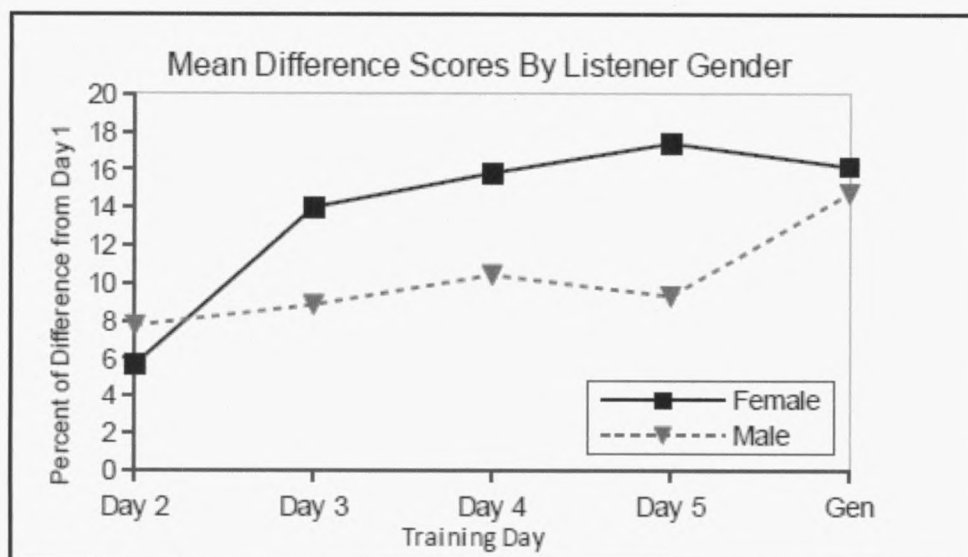
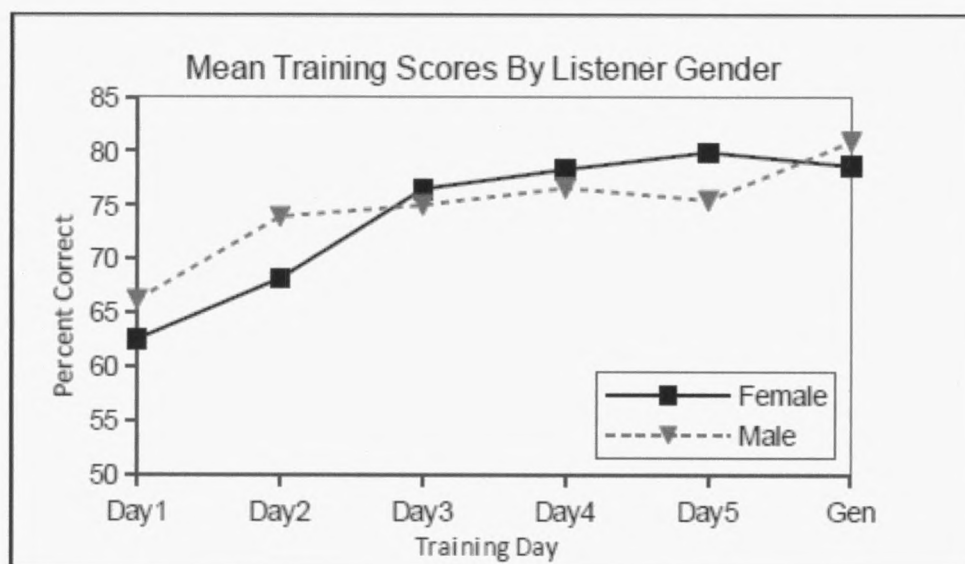
	<i>High Frequency</i>	<i>Low Frequency</i>
Recognition & North vs. South	Become Beyond Market Matter Music Order Person Picture Public System Value Water	Active Beacon Bicep Coffee Dozen Favor Forget Garter Novel Nugget Portal Wedlock
Familiarization	Country Final Police Social	Cousin Listen Nectar Venom
Test Day 1	Basis Between Father Power	Cavern Gusto Handle Symbol
Test Day 2	Common Program Recent Second	Careful Forage Hazel Permit
Test Day 3	Before Better Figure Table	Balance Mingle Staple Title
Test Day 4	Never People Report River	Career Fashion Patron Pigeon
Test Day 5	City Later Single Spirit	Deport Garden Master Rustic
Generalization	Nature Number Party Rather	Captain Jelly Parcel Vision

Figure Captions

Figure 1. Top panel displays the mean scores across training and Generalization by gender of the listener, showing the pattern of learning which occurred for males and females. Bottom panel displays the percentage of change from initial Day 1 scores over training. Females display slightly larger growth over Days 3-5, though male Generalization scores close this gap. Analysis revealed these differences by gender to be of marginal significance.

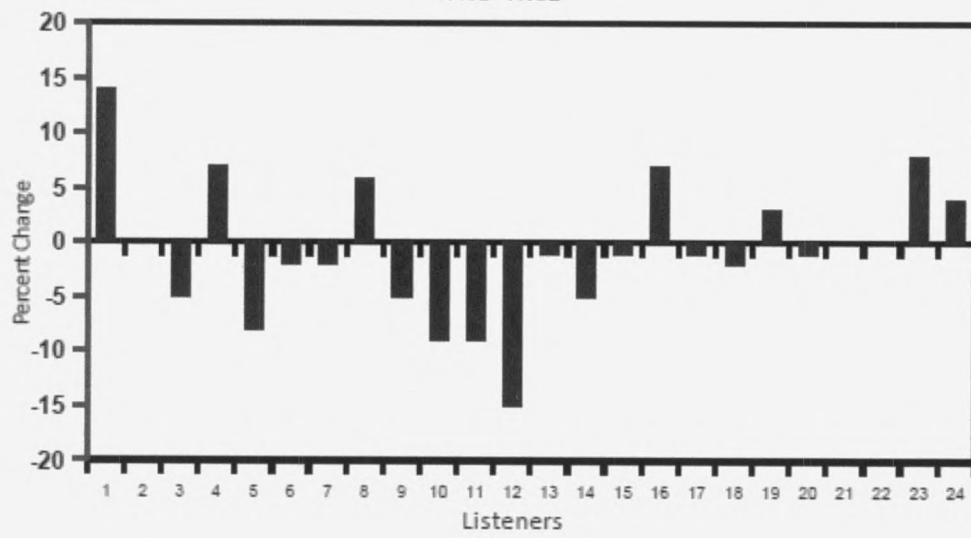
Figure 2. Illustrates the differences in scores from the first dialect discrimination task to the second on an individual basis. Each bar represents a listener, with female comprising the first 12 and males the latter 12. Females displayed a greater overall tendency to decrease in score from the first task.

Figure 3. Displays the change in correctly identifying each word used in the dialect discrimination portion of the study from the first task to the second. Each word is broken down in terms of correct identification when spoken by Northern talkers and Southern talkers. No consistent or significant pattern of identification emerges across words, with the majority of words displaying changes in opposite directions depending on the region of talker.



Difference in Dialect ID Tasks

NvS2 - NvS1



05/01/01

Difference in Correct Identification by Word

